Learning to Estimate: Bayesian Filtering with Deep Density Methods

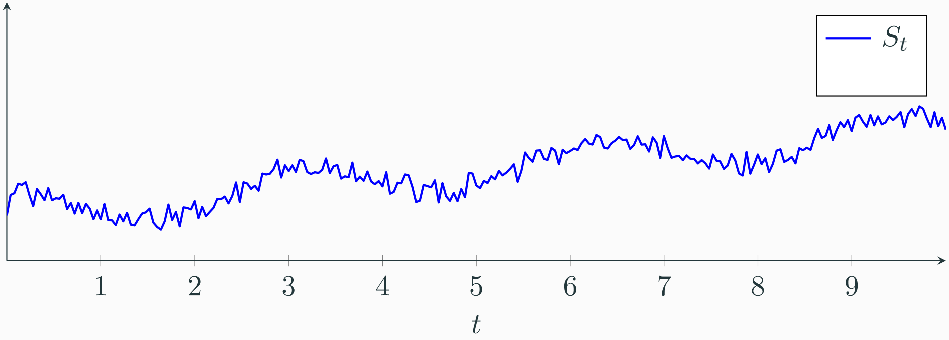
PhD thesis defence, Kasper Bågmark

CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG

Learning to Estimate

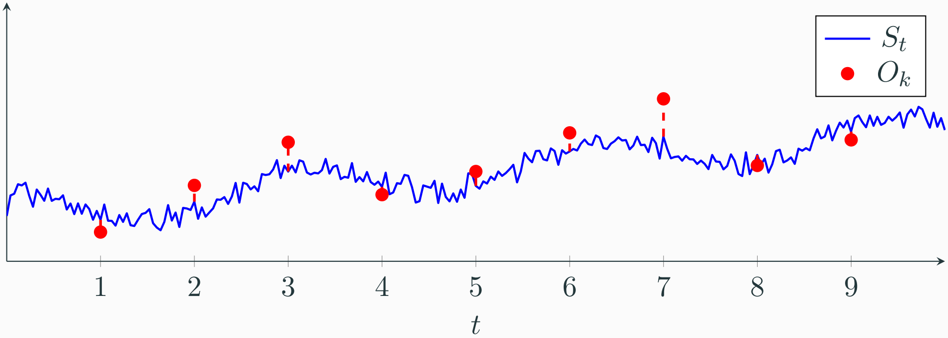
A PhD Thesis in Four Acts

Act One - The Problem



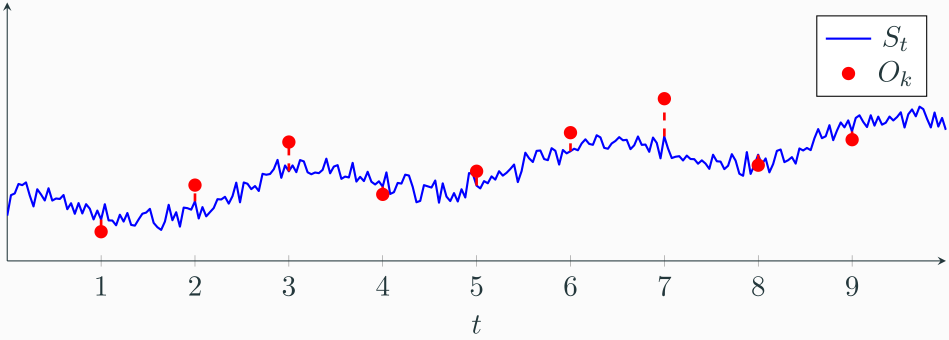
State-space model

$$S_t = S_0 + \int_0^t b(S_u)\,\mathrm{d}u + \int_0^t \sigma(S_u)\,\mathrm{d}B_u,$$



State-space model

$$S_t = S_0 + \int_0^t b(S_u)\,\mathrm{d}u + \int_0^t \sigma(S_u)\,\mathrm{d}B_u,$$
$$O_k = h(S_{t_k}) + V_k$$



Quantity of interest

$$\mathbb{P}\big(S_{t_k} \mid O_{1:k}\big)$$

State-space model

$$S_t = S_0 + \int_0^t b(S_u)\,\mathrm{d}u + \int_0^t \sigma(S_u)\,\mathrm{d}B_u,$$
$$O_k = h(S_{t_k}) + V_k$$

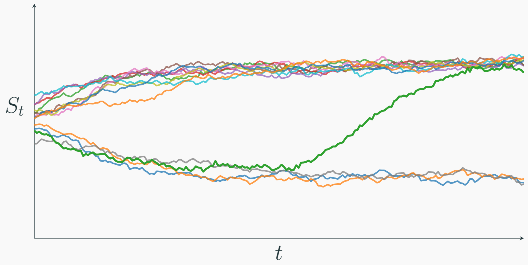
SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

Sample trajectories



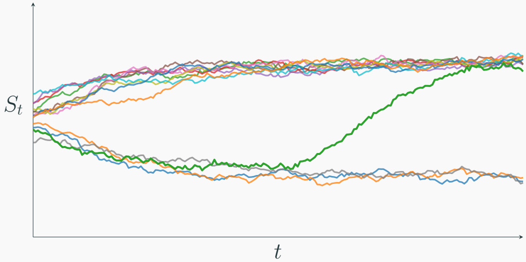
SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

Generator

$$A\phi = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \mu_i \frac{\partial \phi}{\partial x_i}$$

Sample trajectories



SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

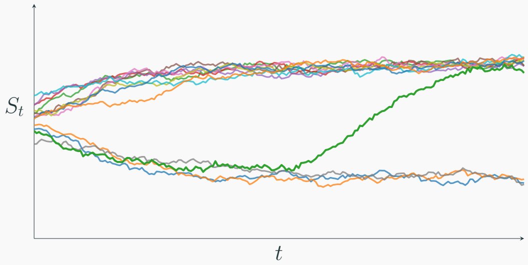
Generator

$$A\phi = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \mu_i \frac{\partial \phi}{\partial x_i}$$

Adjoint

$$A^*\phi = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\left(a_{ij}\phi\right) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(\mu_i\phi\right)$$

Sample trajectories



SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

Fokker–Planck Equation

$$p(t) = p_0 + \int_0^t A^* p(s)\,\mathrm{d}s$$

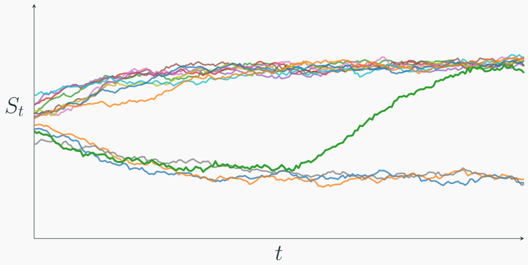
Generator

$$A\phi = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \mu_i \frac{\partial \phi}{\partial x_i}$$

Adjoint

$$A^*\phi = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\left(a_{ij}\phi\right) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(\mu_i \phi\right)$$

Sample trajectories



SDE

$$S_t = S_0 + \int_0^t \mu(S_r)\,\mathrm{d}r + \int_0^t \sigma(S_r)\,\mathrm{d}B_r$$

Fokker–Planck Equation

$$p(t) = p_0 + \int_0^t A^* p(s)\,\mathrm{d}s$$

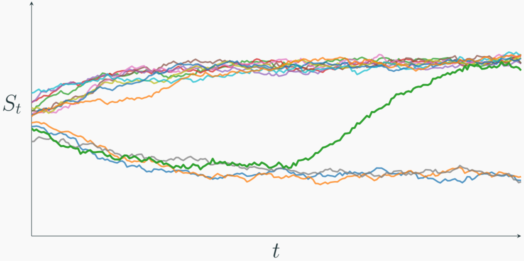
Generator

$$A\phi = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \mu_i \frac{\partial \phi}{\partial x_i}$$

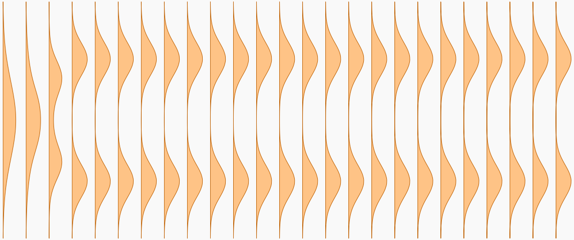
Adjoint

$$A^*\phi = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}\phi) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(\mu_i \phi)$$

Sample trajectories



Density evolution



Target $p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction) $p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\,\mathrm{d}z}$

Target $\quad p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $\quad p_0(0, x) = p(S_0 = x)$

(Prediction) $\quad p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\, \mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $\quad p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$

Prior $p_0(0, x)$ → Prediction $t_0 \to t_1$ → Update o_1 → Prediction $t_1 \to t_2$ → Update o_2 → $\cdots$

Target $\quad p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $\quad p_0(0, x) = p(S_0 = x)$

(Prediction) $\quad p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $\quad p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\,\mathrm{d}z}$

Target $\quad p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $\quad p_0(0, x) = p(S_0 = x)$

(Prediction) $\quad p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\, \mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

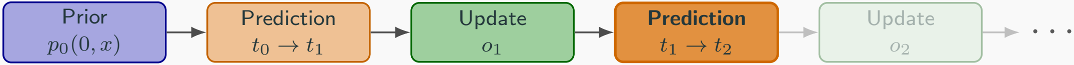
(Update) $\quad p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$

Target $p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction) $p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\, \mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$



Target $\quad p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

$$\text{(Prior)} \quad p_0(0, x) = p(S_0 = x)$$

$$\text{(Prediction)} \quad p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k}) \, \mathrm{d}s, \quad t \in [t_k, t_{k+1}]$$

$$\text{(Update)} \quad p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \frac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z) \, \mathrm{d}z}$$

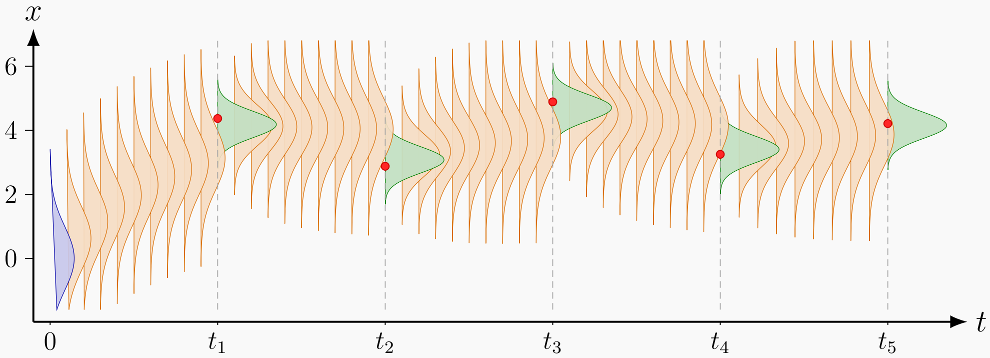
Prior $p_0(0, x)$ → Prediction $t_0 \to t_1$ → Update o_1 → Prediction $t_1 \to t_2$ → **Update** o_2 → $\cdots$

Target $\quad p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $\quad p_0(0, x) = p(S_0 = x)$

(Prediction) $\quad p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $\quad p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\,\mathrm{d}z}$



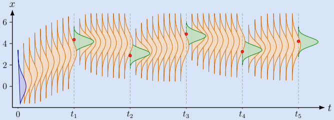
Act One – The Problem

Target $p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction) $p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $p_{k+1}(t_{k+1}, x, o_{1:k+1}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\,\mathrm{d}z}$



o Estimate

A PhD Thesis in Four Acts

Act Two - The Split

Splitting the filtering equation

Operator splitting: **Fokker–Planck**

Splitting the filtering equation

Operator splitting: **Fokker–Planck**

$A^* p = Ap + f(p, \nabla p)$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$$A^*p = Ap + f(p, \nabla p) = Ap + \sum_{i=1}^{d} \left(\sum_{j=1}^{d} \frac{\partial a_{ij}}{\partial x_j} \right) \frac{\partial p}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}\, p - \sum_{i=1}^{d} \frac{\partial \mu_i}{\partial x_i}\, p - 2 \sum_{i=1}^{d} \mu_i \frac{\partial p}{\partial x_i}$$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$A^*p = Ap + f(p, \nabla p) = Ap + f_0\, p + f_1 \cdot \nabla p$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$A^* p = Ap + f(p, \nabla p)$

$$p_t = p_0 + \int_0^t A p_s \,\mathrm{d}s + \int_0^t f(p_s, \nabla p_s) \,\mathrm{d}s$$

Splitting the filtering equation

Operator splitting: **Fokker–Planck**

$A^*p = Ap + f(p, \nabla p)$

$$p_t = p_0 + \int_0^t Ap_s \,\mathrm{d}s + \int_0^t f(p_s, \nabla p_s) \,\mathrm{d}s$$

First order part

$$p_t^{(1)} = p_{t_n}^{(2)} + \int_{t_n}^t f(p_s^{(1)}, \nabla p_s^{(1)}) \,\mathrm{d}s$$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$$A^*p = Ap + f(p, \nabla p)$$

$$p_t = p_0 + \int_0^t Ap_s \,\mathrm{d}s + \int_0^t f(p_s, \nabla p_s) \,\mathrm{d}s$$

First order part

$$p_t^{(1)} = p_{t_n}^{(2)} + \int_{t_n}^t f(p_s^{(1)}, \nabla p_s^{(1)}) \,\mathrm{d}s$$

Second order part

$$p_t^{(2)} = p_{t_{n+1}}^{(1)} + \int_{t_n}^t Ap_s^{(2)} \,\mathrm{d}s$$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$$A^* p = Ap + f(p, \nabla p)$$

$$p_t = p_0 + \int_0^t A p_s \, \mathrm{d}s + \int_0^t f(p_s, \nabla p_s) \, \mathrm{d}s$$

First order part

$$p_t^{(1)} = p_{t_n}^{(2)} + \int_{t_n}^t f(p_s^{(1)}, \nabla p_s^{(1)}) \, \mathrm{d}s$$

Second order part

$$p_t^{(2)} = p_{t_{n+1}}^{(1)} + \int_{t_n}^t A p_s^{(2)} \, \mathrm{d}s$$

After each interval: $p_{t_n}^{(2)} \approx p_{t_n}$, for $n = 0, \dots, N$.

Splitting the filtering equation

Operator splitting: Fokker–Planck

$$A^* p = Ap + f(p, \nabla p)$$

$$p_t = p_0 + \int_0^t A p_s \,\mathrm{d}s + \int_0^t f(p_s, \nabla p_s)\,\mathrm{d}s$$

First order part

$$p_t^{(1)} = p_{t_n}^{(2)} + \int_{t_n}^t f(p_s^{(1)}, \nabla p_s^{(1)})\,\mathrm{d}s$$

$$\downarrow$$

Forward Euler

$$\widehat{p}_{t_{n+1}}^{(1)} = \widehat{p}_{t_n}^{(2)} + \tau f(p_{t_n}^{(2)}, \nabla p_{t_n}^{(2)})$$

Second order part

$$p_t^{(2)} = p_{t_{n+1}}^{(1)} + \int_{t_n}^t A p_s^{(2)}\,\mathrm{d}s$$

Splitting the filtering equation

Operator splitting: Fokker–Planck

$$A^*p = Ap + f(p, \nabla p)$$

$$p_t = p_0 + \int_0^t Ap_s \,\mathrm{d}s + \int_0^t f(p_s, \nabla p_s)\,\mathrm{d}s$$

First order part

$$p_t^{(1)} = p_{t_n}^{(2)} + \int_{t_n}^t f(p_s^{(1)}, \nabla p_s^{(1)})\,\mathrm{d}s$$

$\downarrow$

Forward Euler

$$\widehat{p}_{t_{n+1}}^{(1)} = \widehat{p}_{t_n}^{(2)} + \tau f(p_{t_n}^{(2)}, \nabla p_{t_n}^{(2)})$$

Second order part

$$p_t^{(2)} = p_{t_{n+1}}^{(1)} + \int_{t_n}^t Ap_s^{(2)}\,\mathrm{d}s$$

$\downarrow$

Feynman–Kac representation

$$\widehat{p}_{t_{n+1}}^{(2)}(x) = \mathbb{E}\Big[\widehat{p}_{t_{n+1}}^{(1)}\big(X_{t_{n+1}}\big)\,\Big|\,X_{t_n} = x\Big]$$

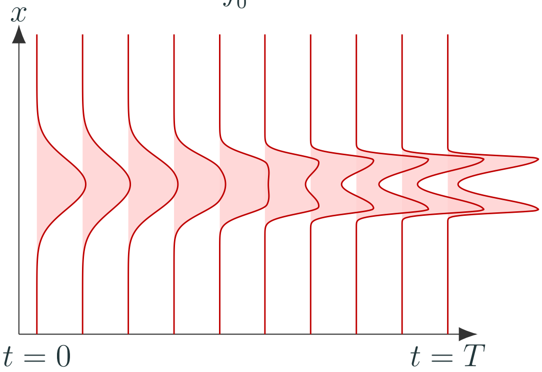
Each split equation is treated with a method adapted to its structure

Splitting the filtering equation

Two split evolutions

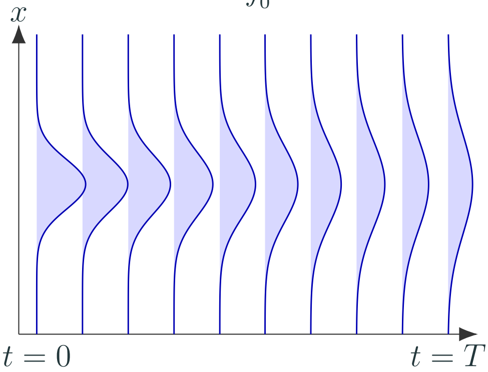
First order part

$$p_t^{(1)} = p_0 + \int_0^t f\left(p_s^{(1)}, \nabla p_s^{(1)}\right) \mathrm{d}s$$



Second order part

$$p_t^{(2)} = p_0 + \int_0^t A p_s^{(2)} \, \mathrm{d}s$$

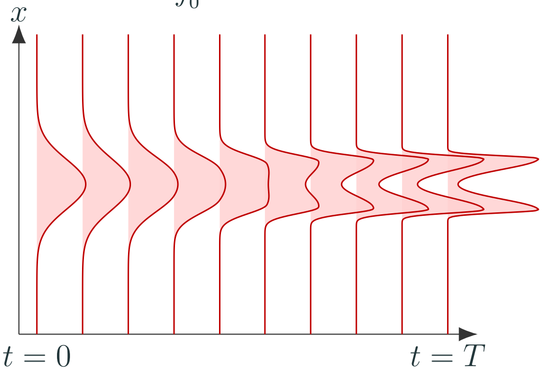


Splitting the filtering equation

Two split evolutions

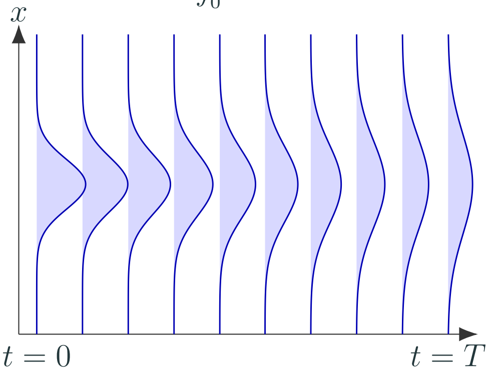
First order part

$$p_t^{(1)} = p_0 + \int_0^t f\left(p_s^{(1)}, \nabla p_s^{(1)}\right) \mathrm{d}s = \Psi_f^t p_0$$



Second order part

$$p_t^{(2)} = p_0 + \int_0^t A p_s^{(2)} \, \mathrm{d}s = \Psi_A^t p_0$$

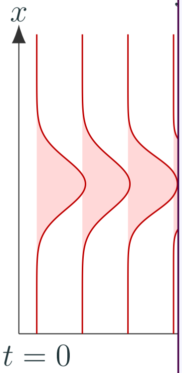


Splitting the filtering equation

Two split evolutions

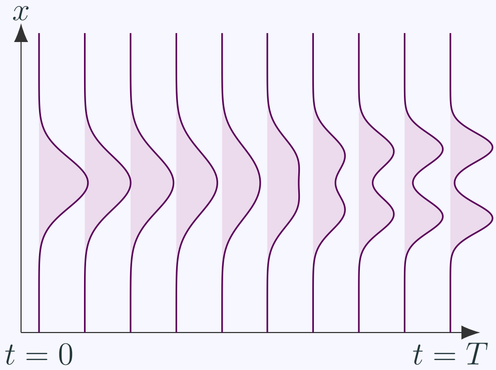
First order pa

$$p_t^{(1)} = p_0 + \ldots$$

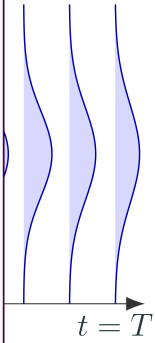


Composed evolution

$$p_{t_n} \approx \left(\Psi_A^{\Delta t} \circ \Psi_f^{\Delta t}\right)^n p_0, \qquad \Delta t = \frac{T}{N}$$



$$\ldots = \Psi_A^t p_0$$



Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Feynman–Kac

$$\widehat{p}^{(2)}_{t_{n+1}}(x) = \mathbb{E}\Big[\widehat{p}^{(1)}_{t_{n+1}}(X_{t_{n+1}}) \,\Big|\, X_{t_n} = x\Big]$$

Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Feynman–Kac

$$\widehat{p}^{(2)}_{t_{n+1}}(x) = \mathbb{E}\Big[\widehat{p}^{(1)}_{t_{n+1}}(X_{t_{n+1}}) \,\Big|\, X_{t_n} = x\Big]$$

Euler–Maruyama

$$\mathcal{X}_{n+1} = \mathcal{X}_n + b(\mathcal{X}_n)\tau + \sigma(\mathcal{X}_n)\Delta W_n$$

Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Feynman–Kac

$$\widehat{p}^{(2)}_{t_{n+1}}(x) = \mathbb{E}\Big[\widehat{p}^{(1)}_{t_{n+1}}(X_{t_{n+1}}) \,\Big|\, X_{t_n} = x\Big]$$

Euler–Maruyama

$$\mathcal{X}_{n+1} = \mathcal{X}_n + b(\mathcal{X}_n)\tau + \sigma(\mathcal{X}_n)\Delta W_n$$

Composed one-step update

$$\overline{\pi}_0(x) = p_0(x)$$

$$\overline{\pi}_{n+1}(x) = \mathbb{E}\Big[\overline{\pi}_n(\mathcal{X}_{n+1}) + \tau f\big(\mathcal{X}_{n+1}, \overline{\pi}_n(\mathcal{X}_{n+1}), \nabla\overline{\pi}_n(\mathcal{X}_{n+1})\big) \,\Big|\, \mathcal{X}_n = x\Big], \quad n = 0, \ldots, N-1$$

Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Feynman–Kac

$$\widehat{p}^{(2)}_{t_{n+1}}(x) = \mathbb{E}\left[\widehat{p}^{(1)}_{t_{n+1}}(X_{t_{n+1}}) \,\middle|\, X_{t_n} = x\right]$$

Euler–Maruyama

$$\mathcal{X}_{n+1} = \mathcal{X}_n + b(\mathcal{X}_n)\tau + \sigma(\mathcal{X}_n)\Delta W_n$$

Composed one-step update

$$\overline{\pi}_0(x) = p_0(x)$$

$$\overline{\pi}_{n+1}(x) = \mathbb{E}\left[\overline{\pi}_n(\mathcal{X}_{n+1}) + \tau f\big(\mathcal{X}_{n+1}, \overline{\pi}_n(\mathcal{X}_{n+1}), \nabla\overline{\pi}_n(\mathcal{X}_{n+1})\big) \,\middle|\, \mathcal{X}_n = x\right], \quad n = 0, \dots, N-1$$

Minimization problem

$$\min_{u \in C(\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|\overline{\pi}_n(\mathcal{X}_{n+1}) + \tau f\big(\mathcal{X}_{n+1}, \overline{\pi}_n(\mathcal{X}_{n+1}), \nabla\overline{\pi}_n(\mathcal{X}_{n+1})\big) - u(\mathcal{X}_n)\right|^2\right]$$

Deep splitting approximation

Compose Forward Euler, Feynman–Kac, and Euler–Maruyama

Forward Euler

$$\widehat{p}^{(1)}_{t_{n+1}} = \widehat{p}^{(2)}_{t_n} + \tau f(p^{(2)}_{t_n}, \nabla p^{(2)}_{t_n})$$

Feynman–Kac

$$\widehat{p}^{(2)}_{t_{n+1}}(x) = \mathbb{E}\Big[\widehat{p}^{(1)}_{t_{n+1}}(X_{t_{n+1}}) \,\Big|\, X_{t_n} = x\Big]$$

Euler–Maruyama

$$\mathcal{X}_{n+1} = \mathcal{X}_n + b(\mathcal{X}_n)\tau + \sigma(\mathcal{X}_n)\Delta W_n$$

Composed one-step update

$$\overline{\pi}_0(x) = p_0(x)$$

$$\overline{\pi}_{n+1}(x) = \mathbb{E}\Big[\overline{\pi}_n(\mathcal{X}_{n+1}) + \tau f\big(\mathcal{X}_{n+1}, \overline{\pi}_n(\mathcal{X}_{n+1}), \nabla\overline{\pi}_n(\mathcal{X}_{n+1})\big) \,\Big|\, \mathcal{X}_n = x\Big], \quad n = 0, \dots, N-1$$

Minimization problem

$$\min_{u \in \mathcal{NN}(\mathbb{R}^d;\mathbb{R})} \mathbb{E}\Big[\Big|\widetilde{\pi}_n(\mathcal{X}_{n+1}) + \tau f\big(\mathcal{X}_{n+1}, \widetilde{\pi}_n(\mathcal{X}_{n+1}), \nabla\widetilde{\pi}_n(\mathcal{X}_{n+1})\big) - u(\mathcal{X}_n)\Big|^2\Big]$$

Convergence result (Paper II)

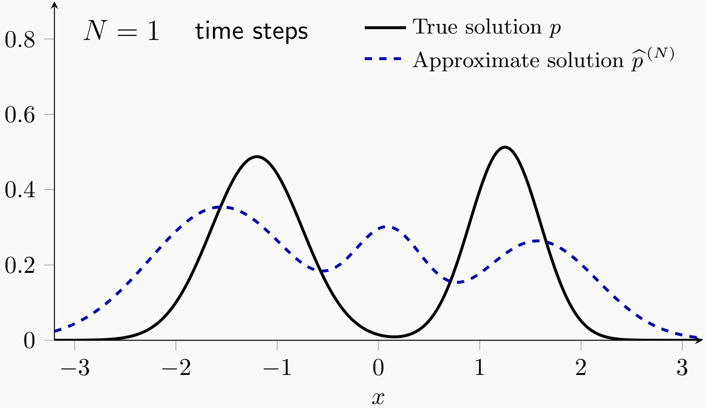
Theorem. Under suitable regularity assumptions, the splitting approximation converges:

$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$

Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

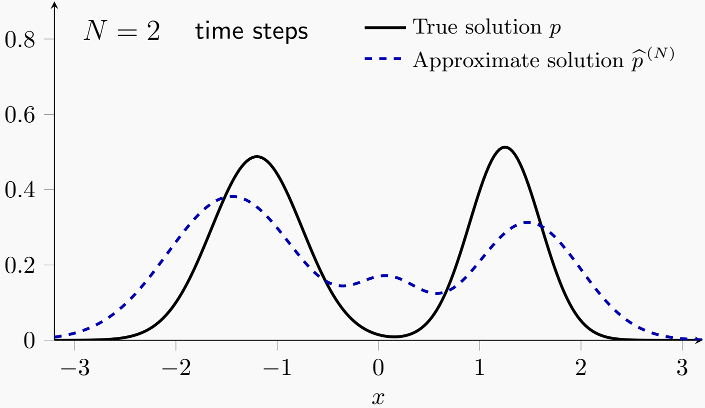
$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq CN^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

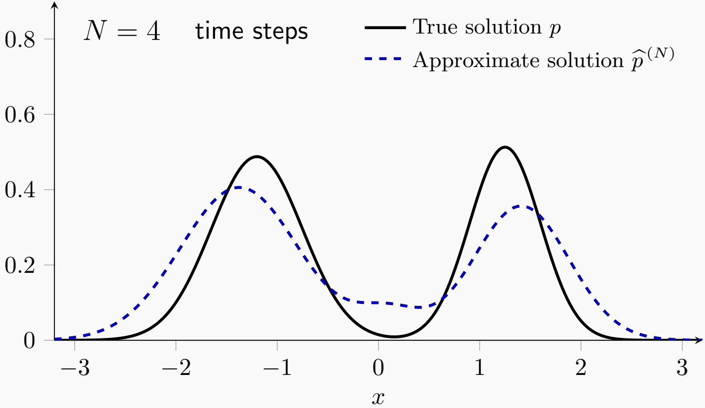
$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq CN^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

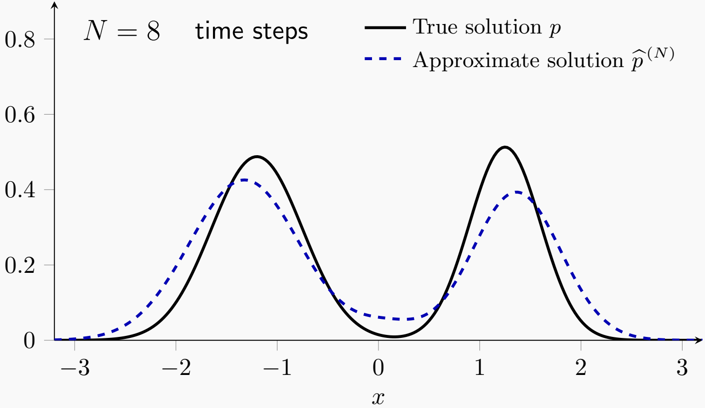
$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}^{(N)}_{t_n} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

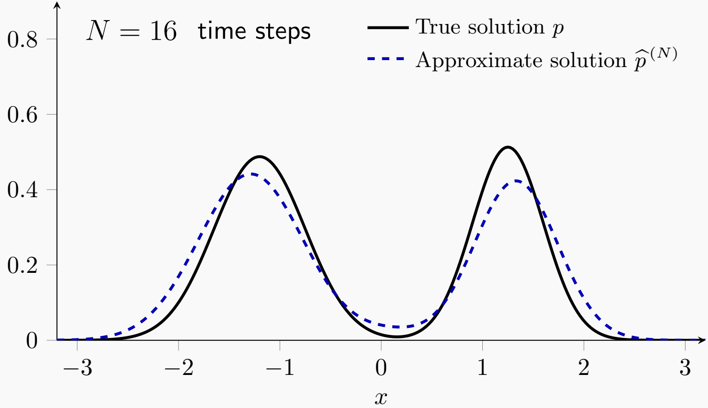
$$\max_{n=0,\dots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \le C N^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

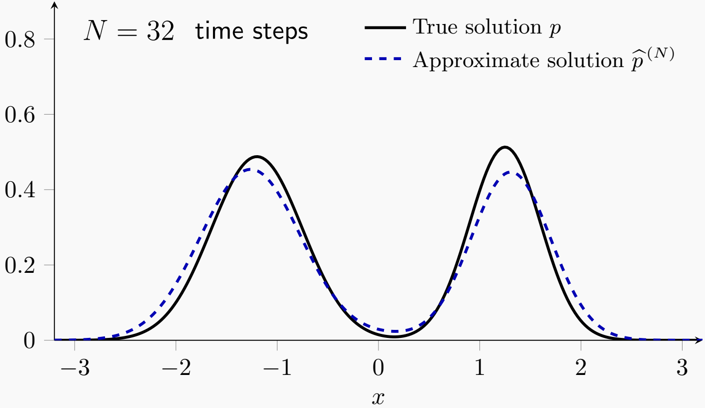
$$\max_{n=0,\dots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

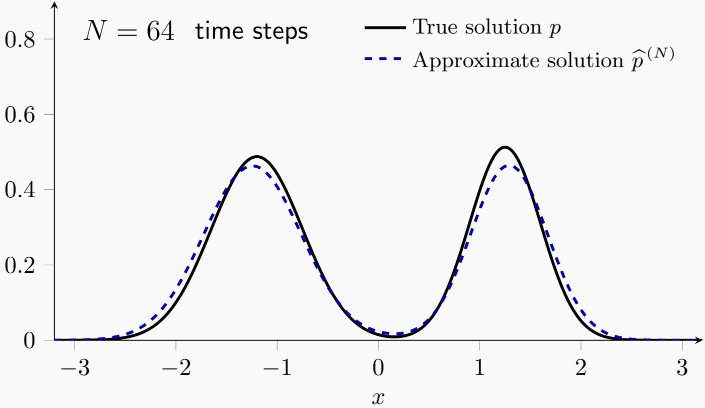
$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}^{(N)}_{t_n} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

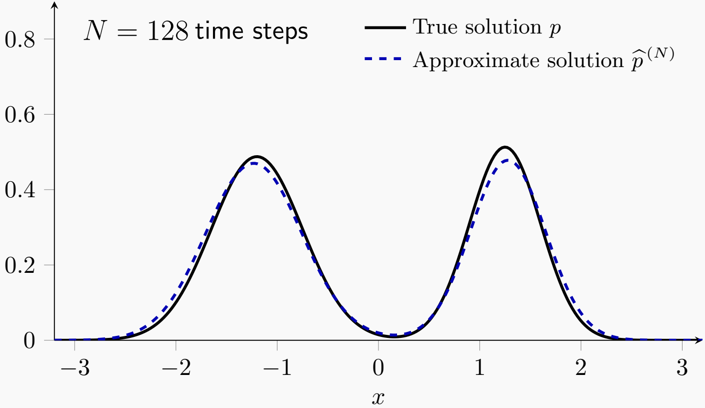
$$\max_{n=0,\dots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

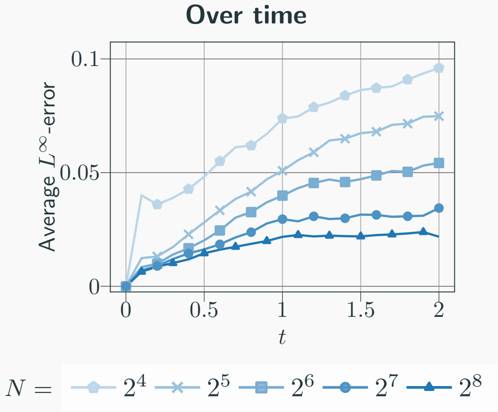
$$\max_{n=0,\dots,N} \left\| p_{t_n} - \widehat{p}^{(N)}_{t_n} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq CN^{-1}$$



Convergence result (Paper II)

Theorem. Under suitable regularity assumptions, the splitting approximation converges:

$$\max_{n=0,\ldots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$

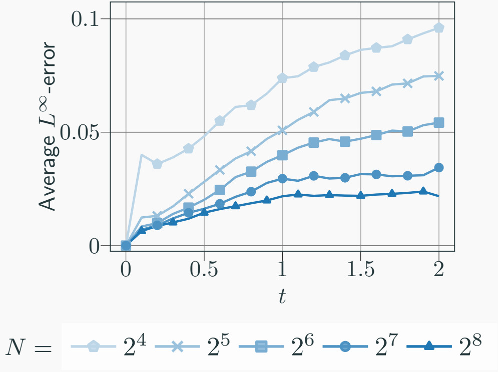


Convergence result (Paper II)

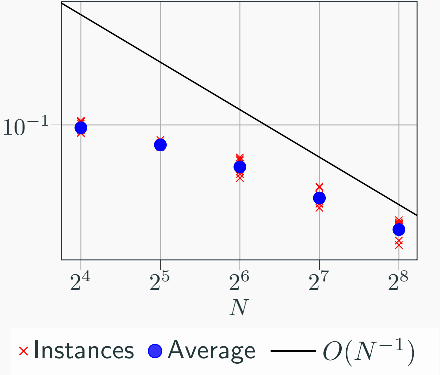
Theorem. Under suitable regularity assumptions, the splitting approximation converges:

$$\max_{n=0,\dots,N} \left\| p_{t_n} - \widehat{p}_{t_n}^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C N^{-1}$$

Over time



Final time



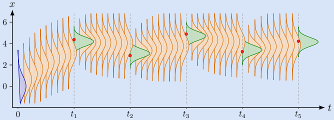
Act One – The Problem

Target $p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction) $p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $p_{k+1}(t_{k+1}, x, o_{1:k+1}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\,\mathrm{d}z}$

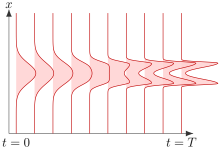


Act Two – The Split

Two split evolutions

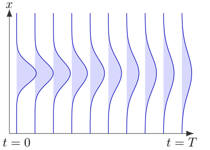
First order part

$$p_t^{(1)} = p_0 + \int_0^t f\left(p_s^{(1)}, \nabla p_s^{(1)}\right) \mathrm{d}s$$



Second order part

$$p_t^{(2)} = p_0 + \int_0^t A p_s^{(2)}\,\mathrm{d}s$$



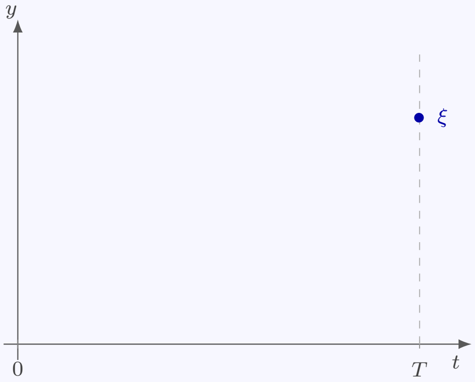
A PhD Thesis in Four Acts

Act Three - The Shooting

Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

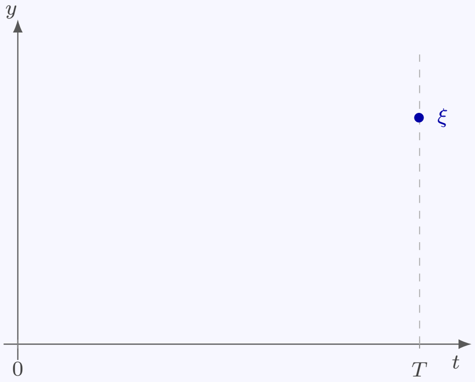


Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

Quantity of interest: find $y(0)$



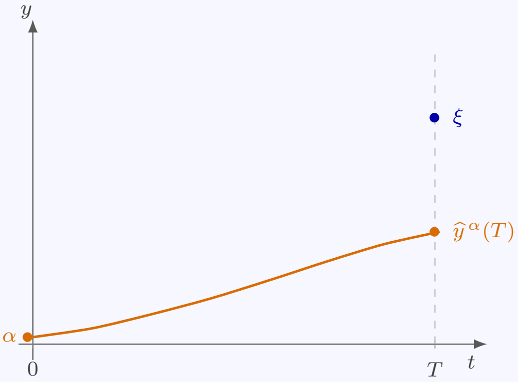
Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward



Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

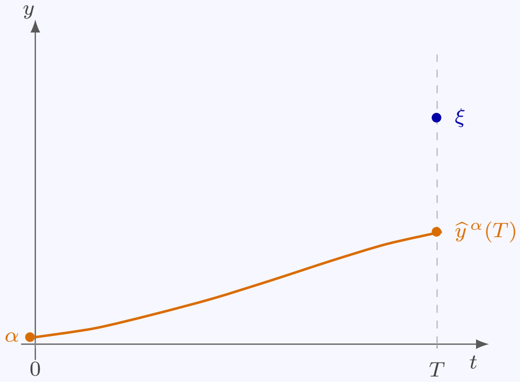
Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward

$$\min_{\alpha} \left| \widehat{y}^{\,\alpha}(T) - \xi \right|$$

Subject to

$$\widehat{y}^{\,\alpha}(t) = \alpha + \int_0^t f(s, \widehat{y}^{\,\alpha}(s))\,\mathrm{d}s, \quad t \in [0, T]$$



Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

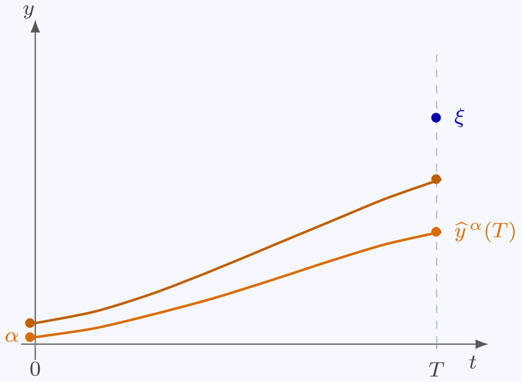
Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward

$$\min_\alpha \left|\widehat{y}^{\,\alpha}(T) - \xi\right|$$

Subject to

$$\widehat{y}^{\,\alpha}(t) = \alpha + \int_0^t f(s, \widehat{y}^{\,\alpha}(s))\,\mathrm{d}s, \quad t \in [0, T]$$



Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

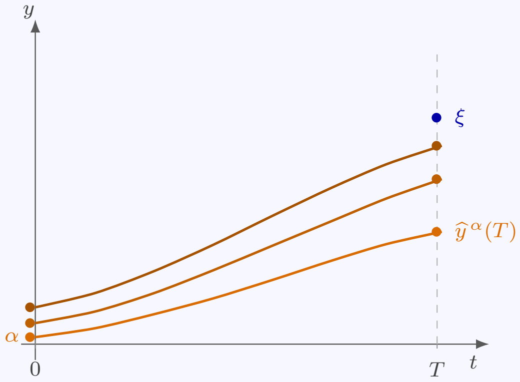
Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward

$$\min_{\alpha} \left|\widehat{y}^{\,\alpha}(T) - \xi\right|$$

Subject to

$$\widehat{y}^{\,\alpha}(t) = \alpha + \int_0^t f(s, \widehat{y}^{\,\alpha}(s))\,\mathrm{d}s, \quad t \in [0, T]$$



Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

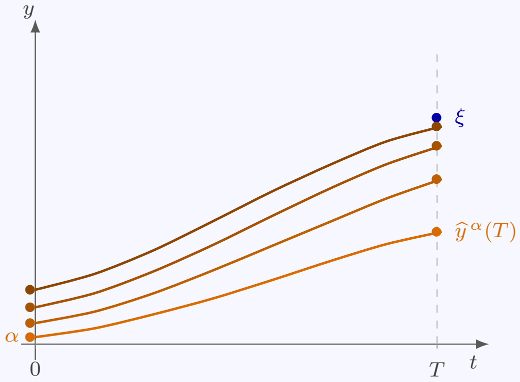
Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward

$$\min_{\alpha} \left| \widehat{y}^{\,\alpha}(T) - \xi \right|$$

Subject to

$$\widehat{y}^{\,\alpha}(t) = \alpha + \int_0^t f(s, \widehat{y}^{\,\alpha}(s))\,\mathrm{d}s, \quad t \in [0, T]$$



Forward shooting

$$y(t) = y(0) + \int_0^t f(s, y(s))\,\mathrm{d}s, \qquad t \in [0, T]$$

$$y(T) = \xi$$

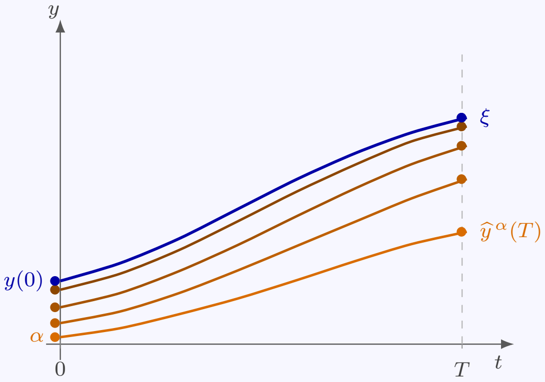
Quantity of interest: find $y(0)$

Guess $\widehat{y}(0) = \alpha$ and integrate forward

$$\min_{\alpha} \left| \widehat{y}^{\,\alpha}(T) - \xi \right|$$

Subject to

$$\widehat{y}^{\,\alpha}(t) = \alpha + \int_0^t f(s, \widehat{y}^{\,\alpha}(s))\,\mathrm{d}s, \quad t \in [0, T]$$



From Fokker–Planck to Backward Stochastic Differential Equation (BSDE)

Continuous-discrete filtering problem

$$
\begin{aligned}
\textbf{(Prior)} \quad & p_0(0, x) = p(S_0 = x) \\
\textbf{(Prediction)} \quad & p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k}) \,\mathrm{d}s, \quad t \in [t_k, t_{k+1}] \\
\textbf{(Update)} \quad & p_{k+1}(t_{k+1}, x, o_{1:k+1}) = \frac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z) \,\mathrm{d}z}
\end{aligned}
$$

From Fokker–Planck to Backward Stochastic Differential Equation (BSDE)

Focus on the prediction step

$$\textbf{(Prior)} \quad p_0(0,x) = p(S_0 = x)$$

$$\textbf{(Prediction)} \quad p(t,x) = p_0(x) + \int_0^t A^* p(s,x)\,\mathrm{d}s, \quad t \in [0,T]$$

$$\textbf{(Update)} \quad p_{k+1} \propto p_k L(o_{k+1}, \cdot)$$

Quantity of interest: $p(T,x)$

From Fokker–Planck to Backward Stochastic Differential Equation (BSDE)

Prediction step

$$p(t,x) = p_0(x) + \int_0^t A^* p(s,x)\,\mathrm{d}s, \qquad t \in [0,T]$$

Nonlinear Feynman–Kac representation

Identification

$$p(T-t, X_t) = Y_t,$$
$$\nabla p(T-t, X_t) = Z_t$$

along the forward diffusion X

BSDE representation

$$Y_t = p_0(X_T) + \int_t^T f\big(X_s, Y_s, Z_s\big)\,\mathrm{d}s - \int_t^T Z_s^\top \sigma(X_s)\,\mathrm{d}W_s, \qquad t \in [0,T]$$

The density value is represented by the backward process Y

From Fokker–Planck to Backward Stochastic Differential Equation (BSDE)

Prediction step

$$p(t,x) = p_0(x) + \int_0^t A^* p(s,x)\,\mathrm{d}s, \qquad t \in [0,T]$$

BSDE representation

$$p(T-t, X_t) = Y_t, \qquad \nabla p(T-t, X_t) = Z_t,$$

$$Y_t = p_0(X_T) + \int_t^T f_s\,\mathrm{d}s - \int_t^T Z_s^\top \sigma_s\,\mathrm{d}W_s$$

Shooting formulation

Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|u(T,X_T) - p_0(X_T)\right|^2\right]$$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\big(X_s, u(s,X_s), \nabla u(s,X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T,x) = u^*(0,x)$

Shooting formulation

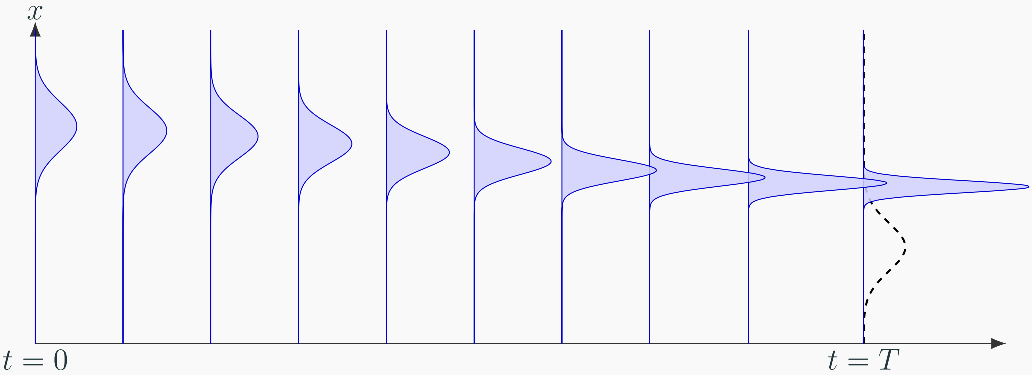
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[|u(T,X_T) - p_0(X_T)|^2\right]$$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\big(X_s, u(s,X_s), \nabla u(s,X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T,x) = u^*(0,x)$



Shooting formulation

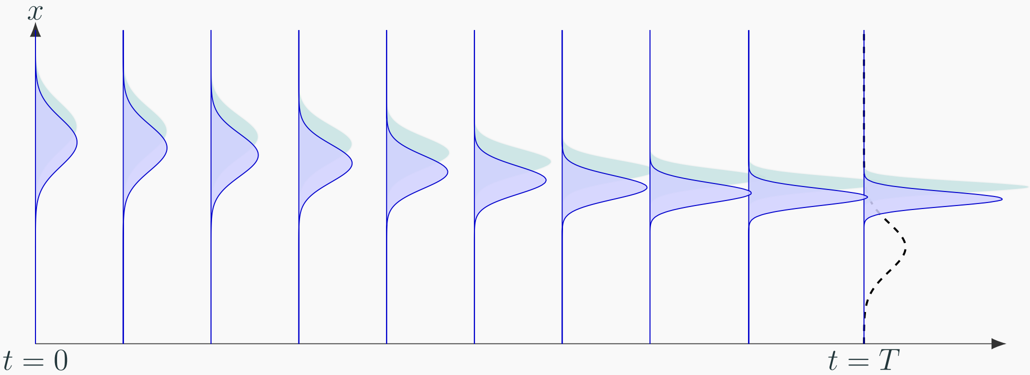
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|u(T,X_T) - p_0(X_T)\right|^2\right]$$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\big(X_s, u(s,X_s), \nabla u(s,X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T,x) = u^*(0,x)$



Shooting formulation

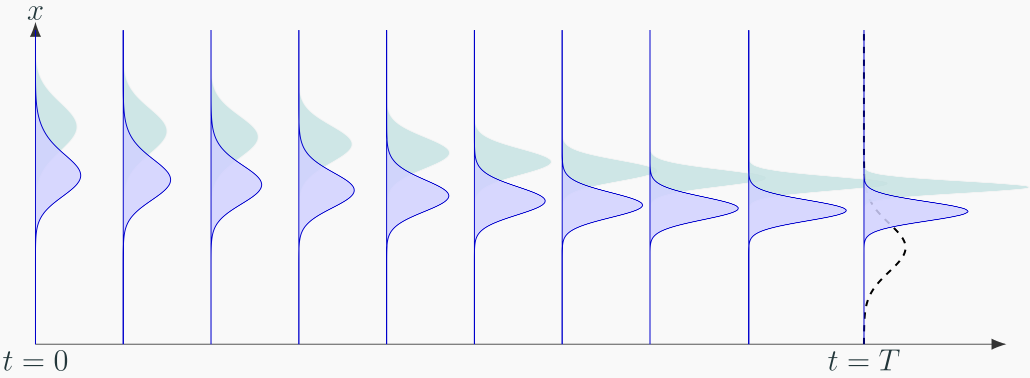
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|u(T,X_T) - p_0(X_T)\right|^2\right]$$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\big(X_s, u(s,X_s), \nabla u(s,X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T,x) = u^*(0,x)$



Shooting formulation

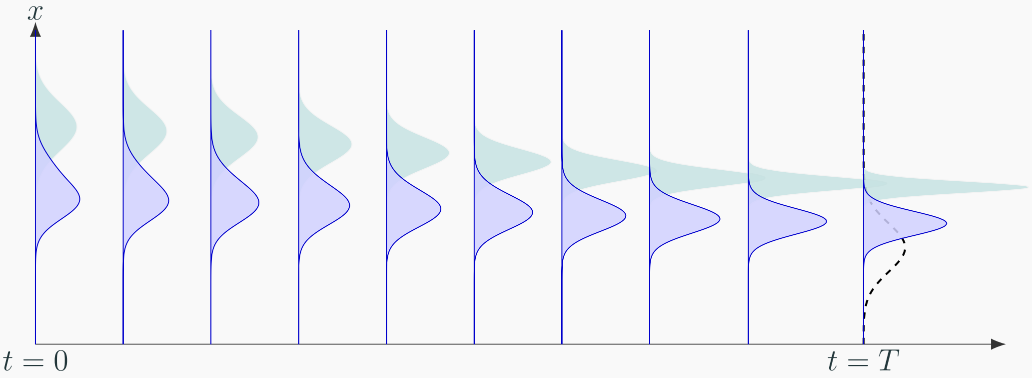
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[|u(T, X_T) - p_0(X_T)|^2 \right]$$

Subject to the dynamics

$$u(t, X_t) = u(0, X_0) - \int_0^t f\big(X_s, u(s, X_s), \nabla u(s, X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s, X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T, x) = u^*(0, x)$



Shooting formulation

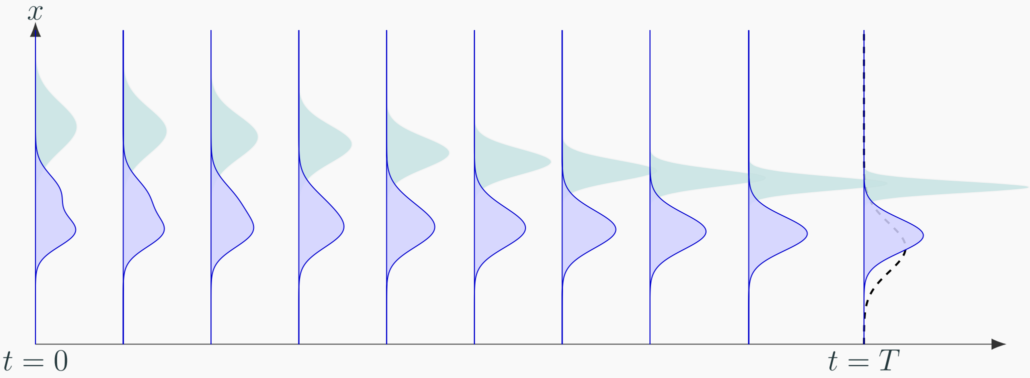
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left| u(T, X_T) - p_0(X_T) \right|^2 \right]$$

Subject to the dynamics

$$u(t, X_t) = u(0, X_0) - \int_0^t f\big(X_s, u(s, X_s), \nabla u(s, X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s, X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T, x) = u^*(0, x)$



Shooting formulation

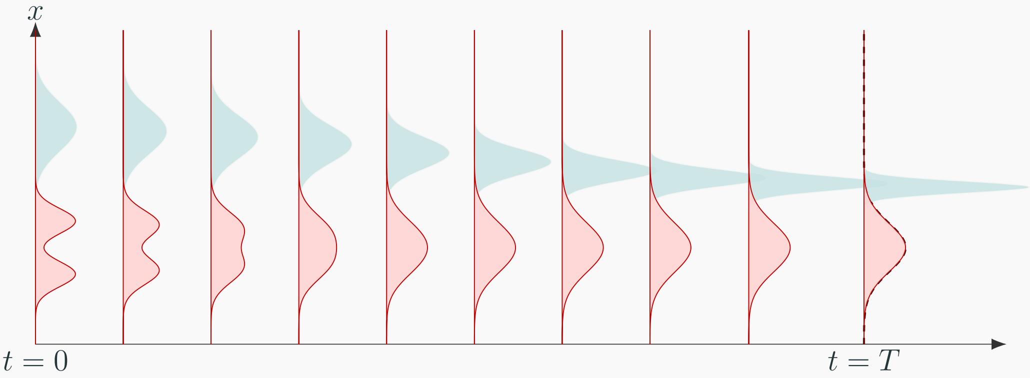
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|u(T,X_T) - p_0(X_T)\right|^2\right]$$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\big(X_s, u(s,X_s), \nabla u(s,X_s)\big)\,\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$

Quantity of interest: $p(T,x) = u^*(0,x)$



Convergence result (Paper III)

Theorem. Under suitable regularity assumptions, and with $\widehat{p}_T = \widehat{u}(0)$, it holds that

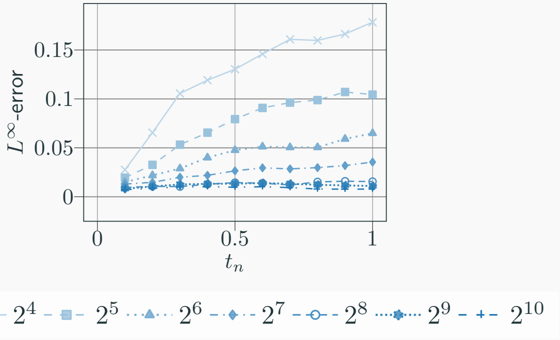
$$\left\| p_T - \widehat{p}_T^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C\left(N^{-\frac{1}{2}} + \sup_x \mathbb{E}\left[\left| \widehat{u}(T, X_T^{0,x}) - p_0(X_T^{0,x}) \right|^2 \right] \right)$$

Convergence result (Paper III)

Theorem. Under suitable regularity assumptions, and with $\widehat{p}_T = \widehat{u}(0)$, it holds that

$$\left\| p_T - \widehat{p}_T^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C\left(N^{-\frac{1}{2}} + \sup_x \mathbb{E}\left[\left| \widehat{u}(T, X_T^{0,x}) - p_0(X_T^{0,x}) \right|^2 \right] \right)$$

Over time

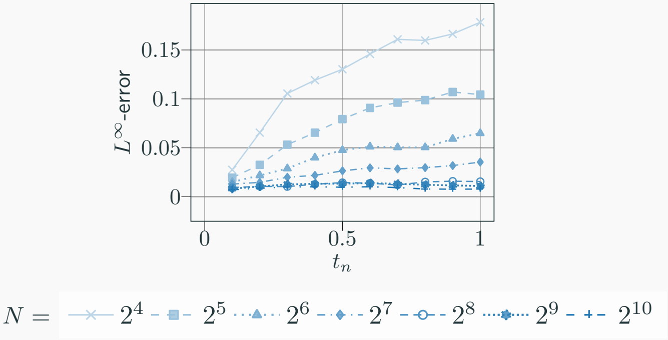


Convergence result (Paper III)

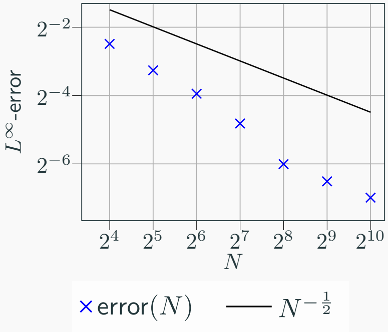
Theorem. Under suitable regularity assumptions, and with $\widehat{p}_T = \widehat{u}(0)$, it holds that

$$\left\| p_T - \widehat{p}_T^{(N)} \right\|_{L^\infty(\mathbb{R}^d;\mathbb{R})} \leq C\left(N^{-\frac{1}{2}} + \sup_x \mathbb{E}\left[\left| \widehat{u}(T, X_T^{0,x}) - p_0(X_T^{0,x}) \right|^2 \right] \right)$$

Over time



Final time



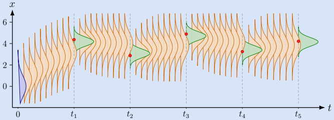
Act One – The Problem

Target $p_k(t,x,o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

$$\text{(Prior)} \quad p_0(0,x) = p(S_0 = x)$$

$$\text{(Prediction)} \quad p_k(t,x,o_{1:k}) = p_k(t_k,x,o_{1:k}) + \int_{t_k}^{t} A^* p_k(s,x,o_{1:k})\,\mathrm{d}s, \quad t \in [t_k, t_{k+1}]$$

$$\text{(Update)} \quad p_{k+1}(t_{k+1},x,o_{1:k+1}) = \frac{p_k(t_{k+1},x,o_{1:k})L(o_{k+1},x)}{\int_{\mathbb{R}^d} p_k(t_{k+1},z,o_{1:k})L(o_{k+1},z)\,\mathrm{d}z}$$

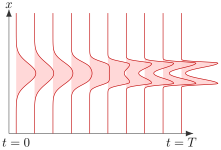


Act Two – The Split

Two split evolutions

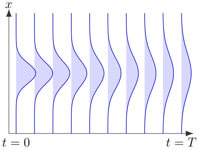
First order part

$$p_t^{(1)} = p_0 + \int_0^t f\left(p_s^{(1)}, \nabla p_s^{(1)}\right)\mathrm{d}s$$



Second order part

$$p_t^{(2)} = p_0 + \int_0^t A p_s^{(2)}\,\mathrm{d}s$$



s in Four Acts

Act Three – The Shooting

Shooting formulation

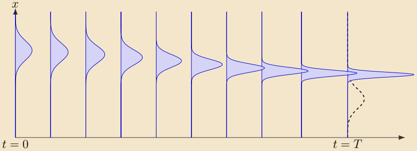
Match terminal condition

$$\min_{u \in C([0,T]\times\mathbb{R}^d;\mathbb{R})} \mathbb{E}\left[\left|u(T,X_T) - p_0(X_T)\right|^2\right]$$

Quantity of interest: $p(T,x) = u^*(0,x)$

Subject to the dynamics

$$u(t,X_t) = u(0,X_0) - \int_0^t f\left(X_s, u(s,X_s), \nabla u(s,X_s)\right)\mathrm{d}s + \int_0^t \nabla u(s,X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$$



Act Four - The Finale

Log-density formulation

From density to log-density

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction)
$$p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} \Big(A p_k(s, x, o_{1:k}) + f\big(x, p_k(s, x, o_{1:k}), \nabla p_k(s, x, o_{1:k})\big)\Big)\, \mathrm{d}s$$
$$t \in [t_k, t_{k+1}]$$

(Update)
$$p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \frac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$$

Log-density formulation

From density to log-density

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction)
$$p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} \Big(A p_k(s, x, o_{1:k}) + f\big(x, p_k(s, x, o_{1:k}), \nabla p_k(s, x, o_{1:k})\big)\Big)\, \mathrm{d}s$$
$$t \in [t_k, t_{k+1}]$$

(Update)
$$p_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = \frac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$$

Define the log-density

$$v_k(t, x, o_{1:k}) = -\log p_k(t, x, o_{1:k}),$$
$$p_k(t, x, o_{1:k}) = \exp\big(-v_k(t, x, o_{1:k})\big).$$

Log-density formulation

From density to log-density

(Prior) $v_0(0,x) = -\log p(S_0 = x)$

(Prediction)
$$p_k(t,x,o_{1:k}) = p_k(t_k,x,o_{1:k}) + \int_{t_k}^{t} \Big(A p_k(s,x,o_{1:k}) + f\big(x, p_k(s,x,o_{1:k}), \nabla p_k(s,x,o_{1:k})\big)\Big)\,\mathrm{d}s$$
$$t \in [t_k, t_{k+1}]$$

(Update)
$$p_{k+1}(t_{k+1},x,o_{1:(k+1)}) = \frac{p_k(t_{k+1},x,o_{1:k}) L(o_{k+1},x)}{\int_{\mathbb{R}^d} p_k(t_{k+1},z,o_{1:k}) L(o_{k+1},z)\,\mathrm{d}z}$$

Define the log-density

$$v_k(t,x,o_{1:k}) = -\log p_k(t,x,o_{1:k}),$$
$$p_k(t,x,o_{1:k}) = \exp\big(-v_k(t,x,o_{1:k})\big).$$

Log-density formulation

From density to log-density

(Prior) $v_0(0,x) = -\log p(S_0 = x)$

(Prediction)
$$v_k(t,x,o_{1:k}) = v_k(t_k,x,o_{1:k}) + \int_{t_k}^{t} \Big(A v_k(s,x,o_{1:k}) + f_{\log}\big(x, v_k(s,x,o_{1:k}), \nabla v_k(s,x,o_{1:k})\big)\Big)\,\mathrm{d}s$$
$$t \in [t_k, t_{k+1}]$$

(Update)
$$p_{k+1}(t_{k+1},x,o_{1:(k+1)}) = \frac{p_k(t_{k+1},x,o_{1:k})L(o_{k+1},x)}{\displaystyle\int_{\mathbb{R}^d} p_k(t_{k+1},z,o_{1:k})L(o_{k+1},z)\,\mathrm{d}z}$$

$$f_{\log}(x,u,w) = -\frac{1}{2}\|\sigma(x)^\top w\|^2 - f(x,1,-w), \qquad x \in \mathbb{R}^d, \quad u \in \mathbb{R}, \quad w \in \mathbb{R}^d.$$

Log-density formulation

From density to log-density

(Prior) $v_0(0, x) = -\log p(S_0 = x)$

(Prediction) $$v_k(t, x, o_{1:k}) = v_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} \Big(A v_k(s, x, o_{1:k}) + f_{\log}\big(x, v_k(s, x, o_{1:k}), \nabla v_k(s, x, o_{1:k})\big) \Big) \, \mathrm{d}s$$

$$t \in [t_k, t_{k+1}]$$

(Update) $$v_{k+1}(t_{k+1}, x, o_{1:(k+1)}) = v_k(t_{k+1}, x, o_{1:k}) - \log\big(L(o_{k+1}, x)\big) + \log\left(\int_{\mathbb{R}^d} \exp\big(-v_k(t_{k+1}, z, o_{1:k})\big) L(o_{k+1}, z) \, \mathrm{d}z \right)$$

$$f_{\log}(x, u, w) = -\frac{1}{2} \|\sigma(x)^\top w\|^2 - f(x, 1, -w), \qquad x \in \mathbb{R}^d, \quad u \in \mathbb{R}, \quad w \in \mathbb{R}^d.$$

Accuracy

Do the methods give accurate estimates?

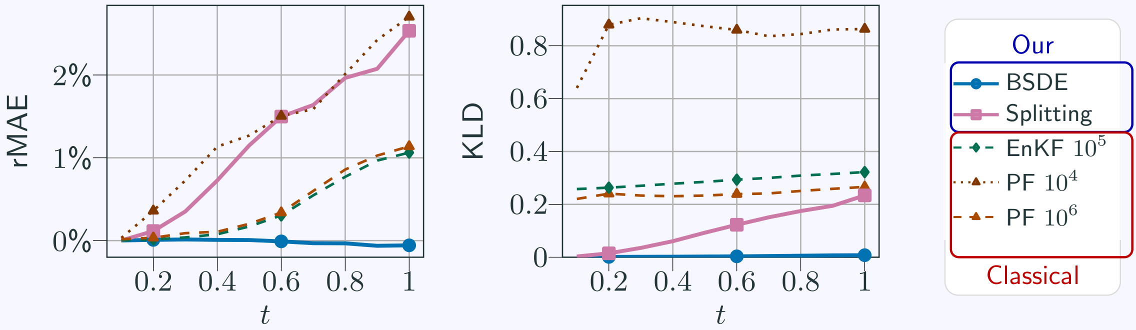
Dimension

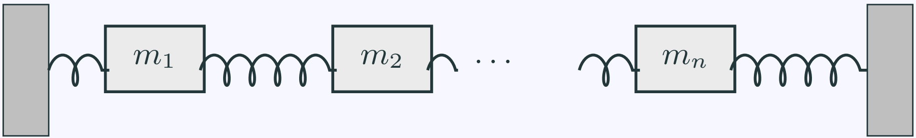
What happens as the dimension increases?

Speed

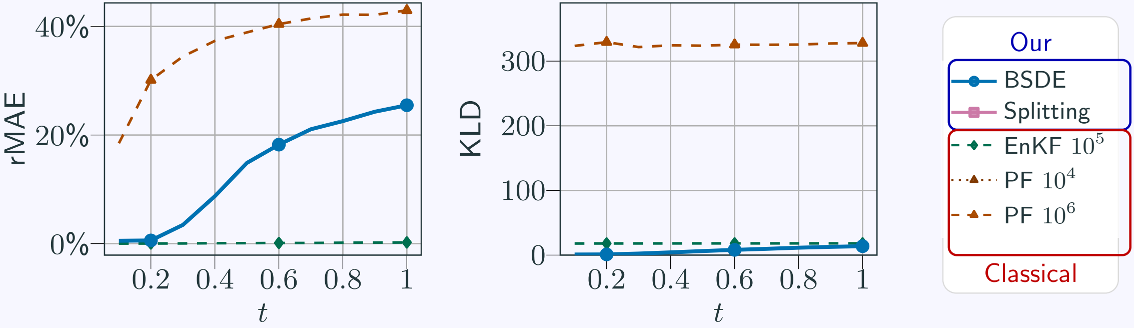
How expensive is inference?

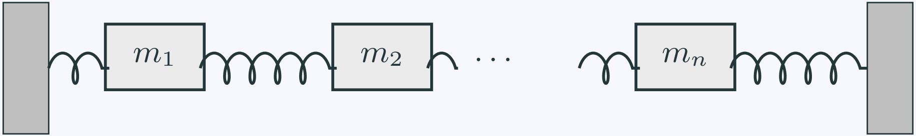
Linear spring–mass 10-dimensional



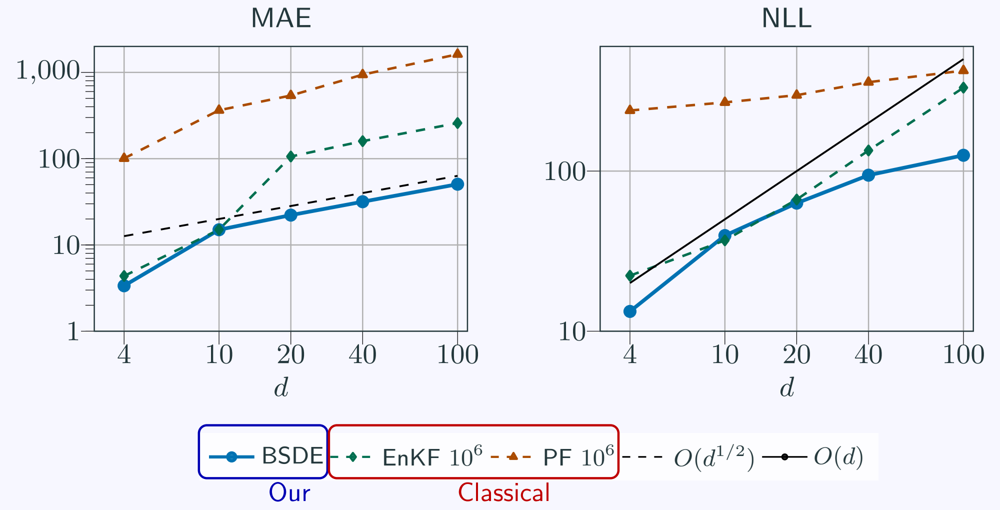


Linear spring–mass 100-dimensional

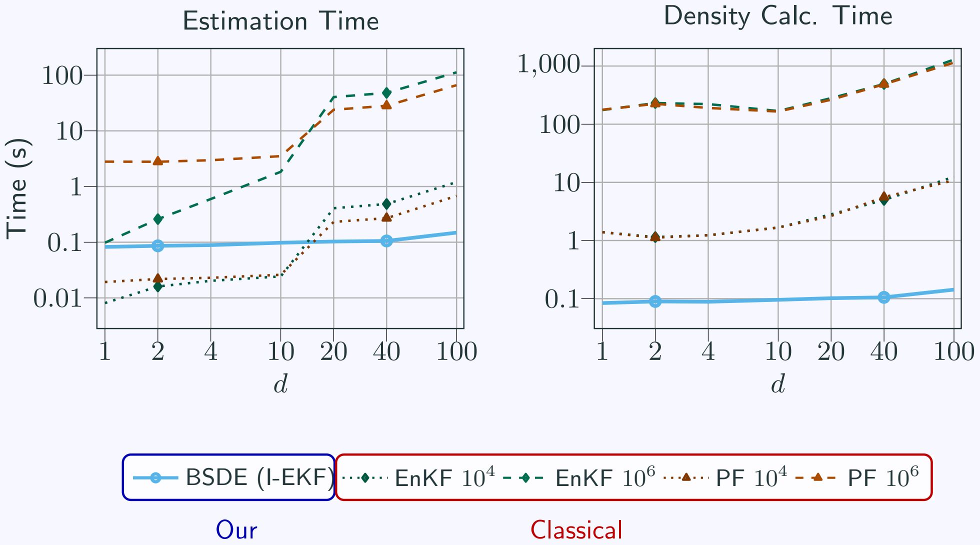




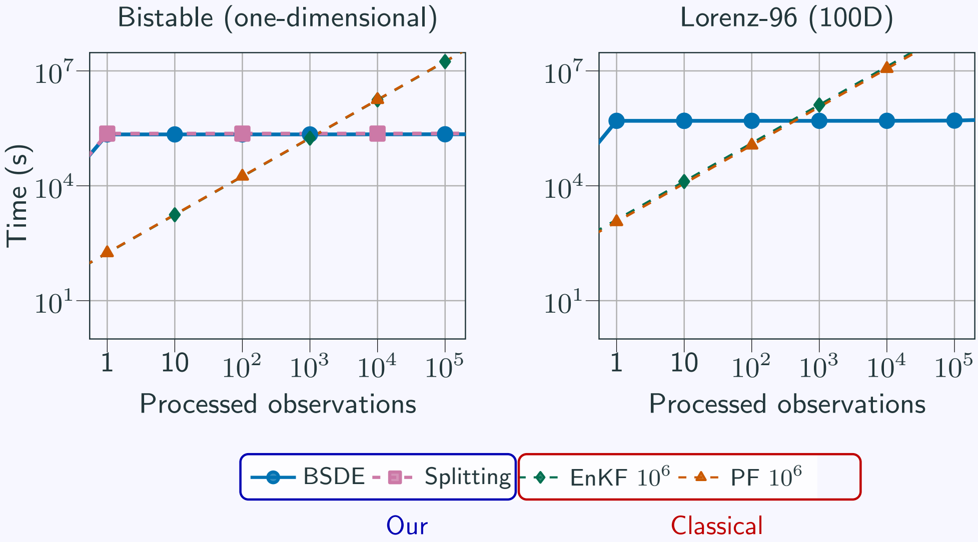
Lorenz–96 (Nonlinear system)



Computational inference time



Trade-off including training time



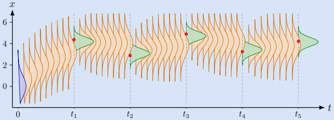
Act One – The Problem

Target $p_k(t, x, o_{1:k}) := p(S_t = x \mid O_{1:k} = o_{1:k})$

(Prior) $p_0(0, x) = p(S_0 = x)$

(Prediction) $p_k(t, x, o_{1:k}) = p_k(t_k, x, o_{1:k}) + \int_{t_k}^{t} A^* p_k(s, x, o_{1:k})\, \mathrm{d}s, \quad t \in [t_k, t_{k+1}]$

(Update) $p_{k+1}(t_{k+1}, x, o_{1:k+1}) = \dfrac{p_k(t_{k+1}, x, o_{1:k}) L(o_{k+1}, x)}{\int_{\mathbb{R}^d} p_k(t_{k+1}, z, o_{1:k}) L(o_{k+1}, z)\, \mathrm{d}z}$

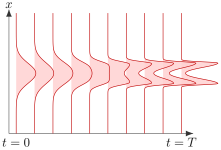


Act Two – The Split

Two split evolutions

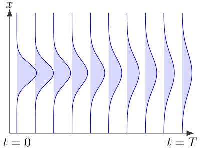
First order part

$$p_t^{(1)} = p_0 + \int_0^t f(p_s^{(1)}, \nabla p_s^{(1)})\, \mathrm{d}s$$



Second order part

$$p_t^{(2)} = p_0 + \int_0^t A p_s^{(2)}\, \mathrm{d}s$$



Act Three – The Shooting

Shooting formulation

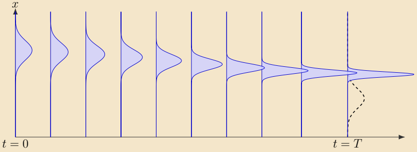
Match terminal condition

$$\min_{u \in C([0,T] \times \mathbb{R}^d; \mathbb{R})} \mathbb{E}\left[|u(T, X_T) - p_0(X_T)|^2 \right]$$

Quantity of interest: $p(T, x) = u^*(0, x)$

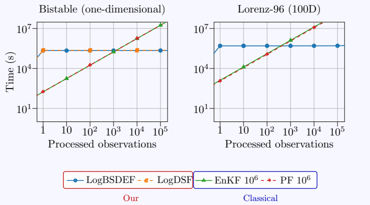
Subject to the dynamics

$$u(t, X_t) = u(0, X_0) - \int_0^t f(X_s, u(s, X_s), \nabla u(s, X_s))\, \mathrm{d}s + \int_0^t \nabla u(s, X_s)^\top \sigma(X_s)\, \mathrm{d}W_s$$



Act Four – The Finale

Trade-off including training time



Conclusion

- Derived simulation-based learning approaches for state estimation
- Proved and numerically verified convergence
- The log-formulation allows for high-dimensional density targets
- Numerically, the BSDE approach outperforms the splitting approach
- The BSDE approach was successful for high-dimensional and nonlinear problems, with favorable computational time compared to classical methods